

Az énekelt magánhangzók fonetikai elemzése

1. Bevezetés. – Beszélőszerveink a mindennapi beszéd létrehozásán kívül számos alternatív beszédtechnikai produkcióra is képesek. Ezek vizsgálata fontos területét képezi a fonetikai kutatásoknak, különösen akkor, ha a mindennapi életben játszott szerepük is jelentős, hiszen produkciójuk és a produktum percepciója is közel mindennapos feladat. Ilyen technika az éneklés is. Fontossága ellenére az éneklés artikulációs, akusztikai és percepciók jellemzőinek vizsgálata még sok szempontból kiaknázatlan terület: bár a hangszínezet és énekhang kapcsolatát (MOLNÁR 1942), illetve az énekhang akusztikai-fizikai jellemzőit (TARNÓCZY 1982) igen, de az éneklés artikulációjának a magyar nyelv beszédhangjaira gyakorolt hatását a jelen cikk szerzőjének kísérletei előtt nem vizsgálták. A következőkben a beszélt és énekelt magánhangzókra és a két csoport közti különbségekre vonatkozó ismereteket foglaljuk össze a nemzetközi szakirodalom alapján, melyek a jelen kutatás kiindulási hipotéziseinek kialakításában is fontos szerepet játszottak.

A magánhangzók akusztikai jellemzői közül meghatározó az első és második formáns (F_1 , F_2), mert ezek mint az alaphang (F_0) legintenzívebb felhangnyalábjai egyértelműen meghatározzák az adott magánhangzót (GÓSY 2004). Bár ez a megállapítás nem az észlelés oldaláról közelít, a következtetés mégis levonható: a hangminőség észlelésében is feltételezhető az alacsonyan lévő, legintenzívebb frekvenciatartomány jelentős szerepe. A következtetés kísérleti úton is bizonyított: egy 2000-ben született kutatás szerint az első két formánst tartalmazó frekvenciasávra szűrt felvételek magánhangzói jó arányban felismerhetőek maradtak (HOLLIEN et al. 2000). Egy magyar nyelvű kísérlet pedig már a második formáns megjelenése előtti alsó frekvenciákra is bizonyította, hogy meglétük elegendő kulcs a magánhangzó-minőség azonosításához (GÓSY 1987).

Artikulációs tekintetben az F_1 értéke elsősorban az állkapocs nyitásszögével, a szájüreg nyitottságával, az ajakműködéssel és a nyelv függőleges mozgásával, míg az F_2 a nyelv alakjával és vízszintes mozgásával van összefüggésben. A magasabb sorszámú formánsok esetében is meghatározhatók ilyen összefüggések: az F_3 értékére szintén a nyelv állása, alakja és az első fogakkal közrezárt tér, az F_4 értékére a vokális traktus gégecső és garat szakasza, míg az F_5 -re a toldalékcső hossza van hatással (LOSCOS 2007: 37). A formánsok létrehozását befolyásoló hangképző szervek és szervszakaszok ismeretében elmondható, hogy a természetes beszéd során leginkább csak az első két-három formáns értékét vagyunk képesek tudatosan változtatni, a többi artikulációs gesztus tudatos irányítása nem vagy csak gyakorlás árán elsajátítható. Mint azt látni fogjuk, az éneklés sajátos artikulációs technikája ezeket a területeket is igénybe veszi.

A magánhangzók észlelését nem csak az akusztikum jellegzetes paraméterei biztosítják – a mindennapi életben felismerésüket megkönnyíti a szemantikailag tartalmas hordozóhangsor, amely a nem tökéletesen artikulált hangok esetében is segíti agyunkat az adott hang felismerésében a beszédpercepciók mechanizmus

visszafelé működő folyamatainak köszönhetően. Ennek megfelelően a csak az ejtési jellegzetességekből fakadó észlelési jellemzők vizsgálatát kizárólag logatomok, azaz értelmetlen hordozóhangsorok használatával biztosíthatjuk. Az értelmetlen hangsorokban az észlelést befolyásolni képes további tényezők a természetes beszédben: a tiszta fázis időtartama, az aktuális hangkörnyezet, a (változatlan formánsszerkezet mellett variált) F0, a formánsok amplitúdója (PETERSON 1961, WENDAHL 1959) és a relatív hangidőtartam (KOVÁCS 1998).

Az énekelt hangzó megkülönböztetése a beszédhangtól auditív úton átlagemberek számára sem igényel különösebb erőfeszítést, ami nem meglepő, hiszen az éneklés artikulációs gesztusai, ebből következően pedig a megváltozott üregeken átjutó zöngé felhangszerkezete (azaz az akusztikum) is jelentősen különbözik a természetes beszédben megfigyelhetőtől.

A beszéd és ének artikulációja közötti legfontosabb különbség a gége süllyesztése, melynek köszönhetően a porcok közötti összekötő szövetek lazábbá válnak. Mivel ezen szövetek és izmok mozgatása szükséges a magas alaphangmagasság eléréséhez, a kevesebb munkával is mozgatható lazább izomállapot lehetővé teszi magasabb F0 képzését is – károsodás nélkül (SUNDBERG 1969, 1970). A gége lejjebb csúsztatásának és a gégeüregek térfogat-növekedésének következtében a gége és a garat kitágul, az így megváltozott toldalékcső pedig (megváltozott sajátfrekvenciája lévén) átalakítja a spektrum magasabb tartományát. Így jön létre az énekesformáns, az énekhang egyik legjelentősebb akusztikai jellemzője, mely az F3-F4-F5 eredeti formánsok helyén egyetlen intenzitáscsúcsként, nagy (férfiaknál 1 kHz-es) sáv szélességen megjelenő felhangnyaláb (SUNDBERG 1972).

A második artikulációs különbség az állkapocs nagyobb szögű nyitottsága, mely a gége lejjebb helyezkedésének fiziológiai kompenzációjaképpen lép fel. Erre válaszként megjelenik az ajkak szélesebbre húzása, az illabiális ejtés is, szintén fiziológiai okból. Mindkét artikulációs gesztus célja az F1 értékének emelése. A negyedik jellemző a toldalékcső kiszélesedése a velum területén, melyhez velárisok esetében a nyelv alakjának változása is társul (SUNDBERG 1969). Ez (a toldalékcső keresztmetszetének növekedésével együtt) a beszélőszervek beállítását az [a] hang ejtési konfigurációjához teszi hasonlóvá. Az így ejtett hangok nagyszámú [a]-ra tévesztésének kimutatásával pedig percepció teszteken keresztül is igazolták a magánhangzóknak a legelső nyelvéllásfokú illabiálishoz történő hasonulását (SCOTTO DI CARLO–GERMAIN 1985).

A módosult artikuláció lehetővé teszi a nagyobb intenzitású hangadást és a magas alaphangok elérését, ám sok esetben a természetes beszéd magánhangzóinak artikulációjával ellentétes tendenciákra kényszeríti a beszélőt, ezzel rontva a produkció és percepció teljesítményszintjét.

Az éneklés artikulációs kényszerének következtében az énekelt magánhangzók akusztikai szerkezete számottevően megváltozik. Az egyik változás a már említett énekesformáns megjelenése, mely jelenség a formánsok kompressziójával jár együtt. A kompressziót az artikulációs gesztusok eredményezik, amelyek következtében a magasabb sorszámú formánsok (F3, F4, F5) is (a természetes beszédben mérhetőnél) alacsonyabb frekvenciaértéken jelennek meg (SUNDBERG 1972).

További különbséget jelent az énekelt és beszélt magánhangzók formánsszerkezetében, hogy az éneklés során az F0 (mely a hangmagasságérzet kialakításáért felelős) a beszédben nem jellemző, magas frekvenciaértéket is felvehet. Ez az F1-re is hatással van: ha az éneklés során gerjesztett F0 a beszédben ejtett vokális átlagos F1 értéke fölé emelkedik, az F1 az artikuláció (az állkapocs és az ajkak működésének) segítségével a megemelt F0 értékére hangolódik, és „összeolvad” vele. A toldalékcso kisélesedése szintén elősegíti ezt a tendenciát (SUNDBERG 1969).

Az akusztikai jellemzők között utolsóként a formánsszűrés, illetve az F0 és F1 értékek emelkedésével magyarázható centralizációs tendenciát említjük. A formánsszűrés változásának következtében a vokális akusztikai szerkezete az F0 emelésével párhuzamosan egyre inkább egymáshoz hasonlóvá válik, míg a 800–1000 Hz-es alaphangmagasság elérésével végül teljes magánhangzóter-redukció jön létre (JOLIVEAU et al. 2004, MILLHOUSE–CLERMONT 2007). Ez azt jelenti, hogy ezen az alaphangmagasságon már nem ejthetők az adott nyelv artikulációs bázisára jellemző magánhangzó-minőségek, csak egyetlen, az énekkhang kiadásához legkényelmesebb fonációs pozíciót biztosító, centralizált helyzetű hang. E hang F1-e az F0 értékével megegyező (~ 1 kHz), F2-je az első felharmonikus (~ 2 kHz) értékére hangolódik. Az előbbiekből következően fontos konklúzióként kiemelendő az is, hogy az alaphangmagasság emelkedésével párhuzamosan egyre kevésbé várható pregnáns magánhangzóejtés, így az ejtési szándéknak megfelelő hang észlelésének lehetősége is csökken.

A felvázolt jellegzetességek tükrében elmondható, hogy magas alaphang képzésekor az éneklés artikulációs kényszere révén mind a produkció, mind a percepció jelentős akadályokba ütközik. A beszélő természetes beszédében mérhető átlagos alaphangmagassághoz közeli énekelt F0 esetén viszont az énekelt vokálisok formánsszerkezete hasonló a beszédben ejtettekéhez (HOLLIEN et al. 2000).

A jelen kutatás célja az énekléstechnika jelenségeinek, illetve a produkció és percepció akadályozottságának vizsgálata a magyar nyelv artikulációs bázisán. Kérdéseinket a nemzetközi (svéd, amerikai, ausztrál, francia és német) szakirodalom eredményeire támaszkodva a következőképpen fogalmaztuk meg. 1. Igaz-e a magyar nyelv magánhangzói esetében is, hogy magas F0 mellett a nagyobb állkapocsnyitás miatt nyitottabb a hangok ejtése, és a percepcióban az ejteni kívánt hangok helyett nyitottabb, alacsonyabb nyelvállásfokú hangok észlelése jellemző? 2. Igaz-e, hogy az illabiális ajakműködés következtében az illabiális hangzók ejtése és észlelése magas F0 mellett is sikeres lehet, míg a labiális hangzókat a hallgatók más, illabiális hangzókként ismerik fel? 3. Igaz-e a magyar nyelv magánhangzói esetében is, hogy a toldalékcso alakjának változása következtében a leggyakoribb az [a:]-ra tévesztés? 4. Létrejön-e a magyar nyelv magánhangzói esetében is a magánhangzóter redukciója, és ha igen, milyen paraméterekkel? 5. Létrejön-e a magyar nyelv magánhangzói esetében is a formánsszűrés?

Hipotézisünk szerint a magyar nyelv artikulációs bázisára elvégzett kísérlet eredményei a nemzetközi szakirodalomban felsorakoztatottakkal összeegyeztethetőek lesznek, de az esetleges eltérések megjelenése nyelvspecifikus sajátosságok létrejöttére utalhat. Elvégzett kísérletünk a kérdéseknek megfelelően több lépésből áll, mely az akusztikum és a percepció vizsgálatát is lehetővé teszi.

2. Anyag, módszer, kísérleti személyek. – A kutatás célkitűzéseinek megfelelően egy három fázisból álló kísérletsorozatot hajtottunk végre. Az első fázis során megadott hangsorok bemondására és feléneklésére kértük az énekest, a második fázisban percepcióos tesztelésnek vetettük alá a rögzített hangsorokat, míg a harmadik fázisban akusztikai elemzéseket végeztünk, különös tekintettel a hangok formánsstruktúrájára. Mivel az éneklésre jellemző hangszínezetet és hangmagasságot a beszélő alapvetően a magánhangzók képzése során alkotja meg, ezért kutatásunk is kizárólag ezen hangok vizsgálatára szorítkozott.

A produkciós fázis hangmintáinak előállítására egy 50 éves koloratúr szoprán hangkategóriájú női szólistát kértünk fel, aki hangját a klasszikus nyugati énektechnikának megfelelően képezte. Mivel a technika egyik legfontosabb jellemzője a magánhangzók hosszú hangoztatásában van, így a kutatás során is csak a hosszú vagy hosszán is ejthető magyar magánhangzókat vizsgálhattuk: [a:], [u:], [i:], [ɔ], [o:], [e:], [ɛ], [ø:], [y:]. A hangmintákat csendesített szobában rögzítettük, a mikrofont az énekes szájától állandó, 20 cm-es távolságban helyeztük el, a felvételeket pedig 44,1 kHz-en, 16 biten digitalizáltuk. 80 db CVC felépítésű logatomot vettünk fel 4 alaphangmagasságon: a beszélő 200 Hz-es beszélt alaphangmagasságán, illetve 500, 550 és 650 Hz-es átlagos énekelt alaphangmagasságon, vibrató nélkül, nem mozgó dallammenettel.

A második, percepcióos fázis során az így előállított logatomokból kiválasztottuk az énektechnikai szakkönyvek szerint a hangképzéshez legelőnyösebb nazális hangkörnyezetet tartalmazókat (KERÉNYI 1959), melyeket auditív megítélés alapján is a legjobb hangminőségűnek tartottunk, ezeket pedig disztraktor hangsorokkal vegyítettük. Így 36 megfigyelt és 194 disztraktor hangsorral folytattuk a kísérletet mind a 4 alaphangmagasságon. A stimulusokat randomizált sorrendben vágtuk össze. Az előzetes kísérletek alapján megállapítottuk milyen hosszúságú szünet ad elegendő időt az adatközlőknek az azonosított hang leírására, a jelen kísérletben pedig az így kimért időtartamoknak megfelelő hosszúságú szüneteket toldottunk a logatomok és bemondott számok közé. Így egy közel 22 perces tesztanyagot készítettünk, melyet 22 darab, darabonként 1 perces blokkra tördeltünk, hogy a blokkok között szükség esetén hosszabb pihenési időt hagyhassunk a tesztelő személyeknek. A vizsgált hangsorok kétszer szerepeltek a kísérlet anyagában. A vágást és összeillesztést az Audacity programmal végeztük.

A hallgatók feladata az volt, hogy a felvétel lejátszásakor lejegyezzék az észlelt magánhangzót egy, a mássalhangzó-környezetet C_C alakban tartalmazó tesztlapon. Ha a tesztelő személy bizonytalan volt a magánhangzó-minőséget illetően, több hangot is megjelölhetett, amelyek között nem tudott egyértelműen dönteni, így egy személy egy stimulusról akár 3-4 adatot is szolgáltatathatott. Emellett lehetőség volt a hangsorok újbóli meghallgatására is. A hallgatók nem vettek részt külön „betanításon”, az énekes hangjával a kísérlet során találkoztak először, de az első blokk eredményeit nem vettük figyelembe, hogy elkerüljük az ismeretlenségből fakadó, esetlegesen rossz eredményeket. Ez a blokk később még egyszer szerepelt a hanganyagban, az erre kapott adatokat már figyelembe vettük.

A percepcióos kísérletben részt vevő 10 személy (4 férfi és 6 nő) átlagos életkora 23 év volt. Az énekelt beszéd észlelésében egységesen átlagos gyakorlottsággal

rendelkeztek, azaz a tesztben nem vettek részt ének- vagy beszédtechnikával foglalkozó személyek.

A kutatás harmadik fázisában a korábban észlelési tesztnek alávetett hangsorok magánhangzóinak formánsértékeit mértük a Praat program segítségével. A méréseket manuálisan, a tiszta fázisnak ítélt szakasz közepén, egy pontban végeztük el. Mivel a nazális környezetben ejtett magánhangzók szegmentálása problémás kérdés, fontos megemlíteni, hogy azokban az esetekben, ahol a hanghatárok nem voltak egyértelműek, illetve ahol folyamatos volt a formánsmozgás, a mérést a formánsátmenet számtani középpontjában végeztük el. A mért adatok alapján megvizsgáltuk a formánskompresszió megjelenését, a magánhangzók centralizálódásának tendenciáját, a magas alaphangon is jól felismerhető vokálisok akusztikai szerkezetének változását, és a következőket tévesztett hangok akusztikai szerkezetének hasonlóságát is. Megmértük a legmagasabb alaphangon jó arányban felismerhető magánhangzó-realizációk időtartamát is, annak vizsgálatára, hogy van-e összefüggés a realizáció időtartama és a beszélő ejtési szándékának megfelelő észlelés között.

3. E r e d m é n y e k . – A percepciók tesztelés során adott válaszokat alaphang-magasságönként tévesztési mátrixokban összesítettük. A beszélt alaphangmagasságon rögzített stimulusok felismerési aránya közel 95%-os volt, mely eredmény biztosítékot jelentett arra nézve, hogy a beszélő artikulációja tiszta. Az ejtési szándéknak megfelelő hangazonosítás aránya a további alaphangmagasságokon csökkent, bár nem az elvárt, lineáris módon: 500 Hz-es F0 esetén 53%, 550 Hz-es F0 esetén 67%, 650 Hz-es F0 esetén 38% volt.

1. ábra

A tévesztési mátrixok: $F_0 = 500$ Hz, $F_0 = 550$ Hz, $F_0 = 650$ Hz

Ejteni szándékozott magánhangzó	a:	ɔ	ø:	y:	ε	e:
a:	5	12	5	1		
ɔ		17	4	1		
ø:		2	13	6		
y:			3	17		
ε				1	15	3
e:					11	9
i:					16	5
o:						5
u:						16
ü:						8
ö:						2

Ejteni szándékozott magánhangzó	Észlelt magánhangzó								
	a	ɔ	o	u	ø	ɤ	ɛ	e	i
a:	20								
ɔ	3	15	3						
o:		5	12	5					
u:		3	9	9					
ø:					11	11			
ɤ:					6	15			
ɛ							1	19	
e:									15
i:						2		8	14

[illegible]

Az első énekelt alaphangmagasságon ($F_0=500$ Hz) nagy szórásban mutatkoztak tévesztések. A veláris csoporton belül a nyíltabbra tévesztés helyett kizárólag zártabb hangzóra tévesztések fordultak elő, a legalsó nyelvállású illabiális [a:] hang felismerési aránya az elvárthoz képest alacsony, 22% volt. A palatálisok között nyíltabbra tévesztés két esetben fordult elő (az [i:] hang esetében), minden egyéb esetben (a veláris csoportnál láthatóhoz hasonlóan) zártabb hangzót észleltek a hallgatók. Ezek között nagy számban jelent meg az [e:]–[i:] tévesztés (76%), illetve az [ø:]–[y:] tévesztés (63%). A labiálisok illabiálisra tévesztése az [y:] és [ø:] hangok esetében is előfordult, az [ø:] esetében 33%-ban, az [y:] esetében 55%-ban érvényesült a tendencia. A legjobb felismerési arányt a zárt illabiális [i:] hang (90%), a veláris csoportban pedig az [u:] hang (85%) mutatta. A legrosszabb arányban felismert hangzó az [ø:] volt (4%).

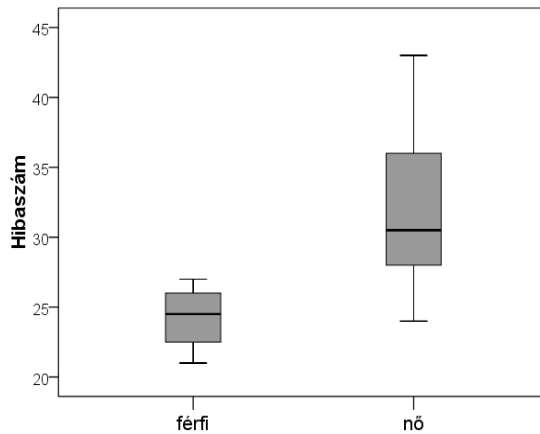
A második énekelt alaphangmagasságon ($F_0=550$ Hz), az elvárttól eltérő módon, az előzőhöz képest jobb arányban sikerült a hallgatóknak az ejteni kívánt hangok pontos azonosítása. Ezen a magasságon két hangzó felismerése maradt a legstabilabb: az alsó nyelvállású illabiális veláris [a:] (100%) és az alsó nyelvállású illabiális palatális [ɛ] (95%) hangoké. A két hang ejtése az éneklés artikulációs kényszerének kedvez, így jó arányú felismerésük az elvárásainknak megfelelő. A nyíltabbra tévesztés tendenciája az előzőekhez mérten számottevően erősebben érvényesül: [ɔ]–[a:] 14%-ban, [o:]–[ɔ] 23%-ban, [u:]–[o:] 43%-ban, [u:]–[ɔ] 14%-ban, [y:]–[ø:] 29%-ban, [i:]–[e:] 33%-ban. A labiális hangzók illabiálisra tévesztése kis számban jelenik meg, és kizárólag a veláris csoportban, az [ɔ] hang esetében: 14%-ban [a:]–ként ismerték fel a hallgatók. A legnagyobb számú következetes tévesztés ismételten az [ø:]–[y:] (50%) és [e:]–[i:] (40%) hangokat érinti.

A legmagasabb alaphangon a válaszok szórása az előzőekhez képest nagyobb, ami azt jelenti, hogy a téves válaszok egyöntetűsége csökkent. A legjobban felismerhető veláris hangzó (nem várt módon) a felső nyelvállású labiális [u:] hang (43%), a palatális csoportban ugyanez a zárt labiális [y:] (65%), a középső nyelvállású illabiális [ɛ] (67%) és a zárt illabiális [i:] (67%) hangokról mondható el. Az [y:] hangzó jó felismerési aránya képzési jegyeiből következően nem várt eredmény, az [i:] hangzó illabialitása révén részben, az [ɛ] pedig nyitottsága és illabialitása révén teljesen kiszolgálja az éneklés artikulációs kényszerűségeit, így ezek pontosabb artikulációját és észlelését elvártuk. A legtöbb következetes tévesztés az [ø:]–[i:] (68%) és [e:]–[i:] (58%) hangpárokat érintette. Mindkét esetben a zártabb hangzót észlelték a tesztelők, bár az [ø]–[i:] pár esetében a labiális hangzó helyett az illabialist jelölték, mely ajakrészessége révén valóban közelebb áll az éneklés által megkívánt artikulációhoz. Ezekhez képest a nyíltabb hangzó észlelésének tendenciája alacsonyabb rátát mutat: a veláris csoporton belül az [u:]–[o:] (14%), az [u:]–[ɔ] (43%) és az [o:]–[ɔ] (15%) párokat, a palatális csoporton belül a [ø:]–[ɛ] (9%), az [y:]–[e:] (21%), az [i:]–[e:] (33%), az [e:]–[ɛ] (4%) párokat érintette. Az illabiális hangzók észlelése az ejteni kívánt labiális helyett szintén kisebb arányban jelentkezik: a veláris csoportban nem találunk [a:] jelölést bármely más veláris helyett, a palatális csoportban a tendencia a már említett [ø:]–[i:] (68%), illetve az [ø:]–[e:] (18%), az [ø:]–[ɛ] (9%), az [y:]–[e:] (22%) és az [y:]–[i:] (13%) párok esetében érvényesült.

A beszédpercepció kísérletben a legnagyobb számú hibás válaszként az [i:] jelent meg, azaz az ejteni kívánt énekelt magánhangzó helyett legnagyobb számban ezt a felső nyelvállású illabiális hangzót jelölték a hallgatók. A hang 27%-ban szerepelt a hibás válaszok között. A velárisok között legnagyobb számban az ajakkerekítéses alsó nyelvállású [ɔ] hangot (16%) jelölték a hallgatók. Az éneklés artikulációs kényszerének legjobban megfelelő ejtési konfigurációjú [a:] hangzóra mindössze a hibás válaszok 1%-ában tévesztettek.

2. ábra

A női és férfi tesztelők tévesztéseinek szóródása



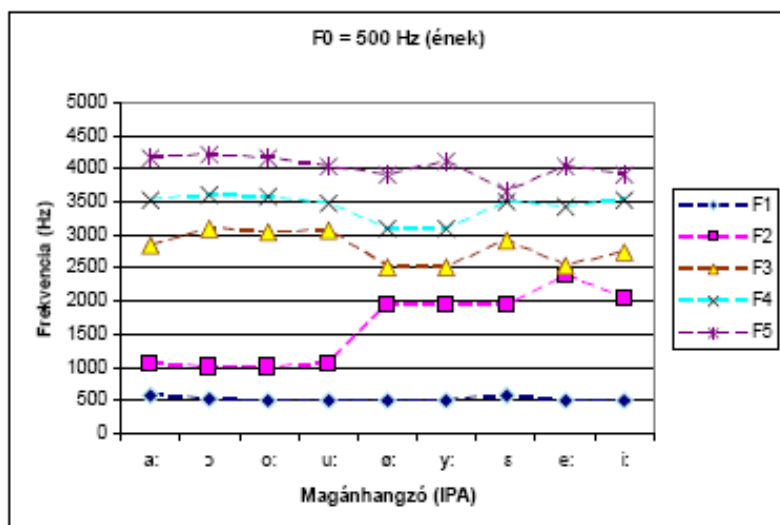
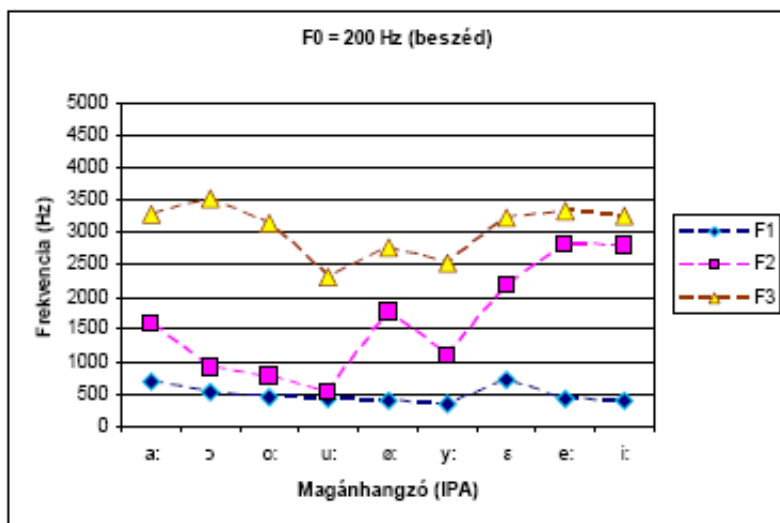
A nők és férfiak eredményeinek megoszlásából azt látjuk, hogy a nők átlagosan több hibát és nagyobb szóródásban produkáltak. A 4 férfi összesen 97-szer, míg a 6 nő összesen 192-szer tévesztett. A nők csoportjában két, az átlaghoz képest kiugróan nagy számú hibás eredményt (36 és 43 hiba) is találunk.

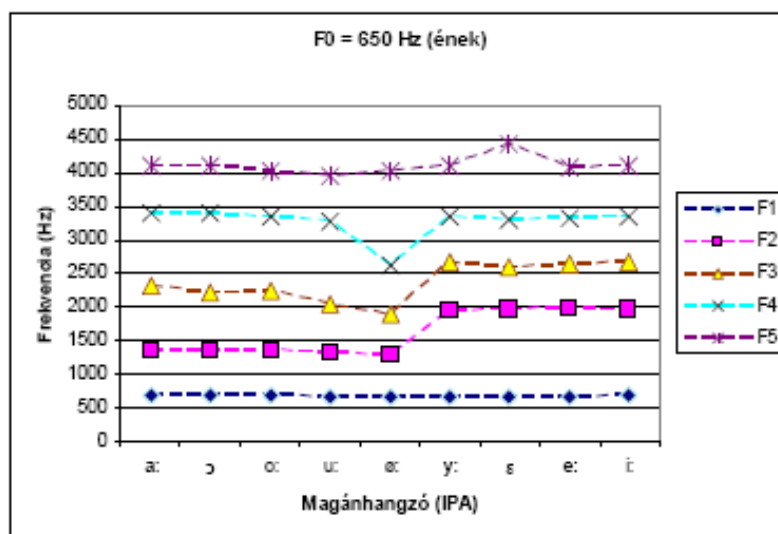
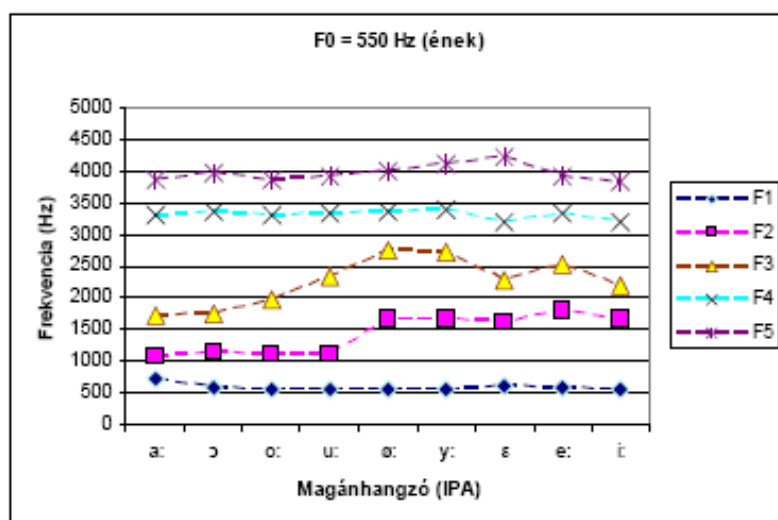
Az akusztikai mérések során megvizsgáltuk 1. a formánskompresszió megjelenését, 2. a magánhangzótér redukcióját és annak mértékét, 3. a legmagasabb alaphangon is az ejtési szándéknak megfelelően jó arányban felismerhető hangok akusztikai szerkezetének változását, illetve 4. a nagy számú következetes hangtévesztésekben szereplő vokálist párokat, különös tekintettel arra, hogy ez utóbbiak az éneklés artikulációs kényszerével ellentmondásban nyíltabbról zártabbra tévesztésben kerültek kapcsolatba.

A formánskompresszió megjelenését és mértékét szemlélteti a 3. ábra.

3. ábra

A magánhangzók formánsértékeinek változása alaphang-magasságonként



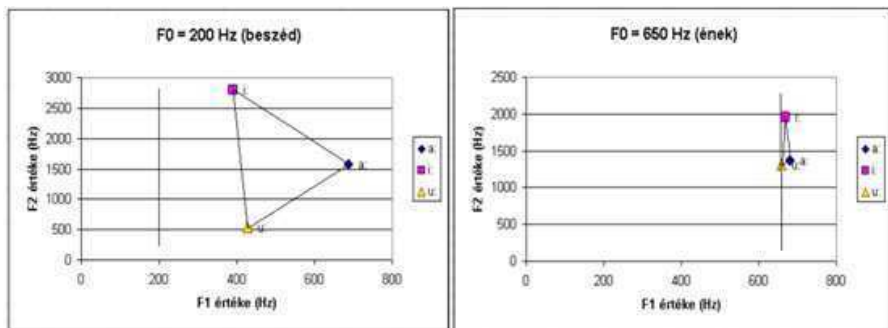


Az ábráról leolvasható, hogy az énekelt vokálisok formánsai a beszédben mérhető frekvenciaértékeihez képest jelentősen alacsonyabb frekvenciaértéken realizálódnak. Főként a magasabb sorszámú formánsok értéke csökken: míg a beszédben mérhető F3 a 2,5–3,5 kHz-es tartományban jelenik meg, addig az éneklés során az alaphangmagasság emelésével egyre szűkebb tartományban, egyre alacsonyabb frekvenciaértéken mértük. Az ábra ezen kívül rámutat az akusztikai szerkezetek egymáshoz történő hasonulására is: a hasonulás alacsonyabb F0 esetén veláris és palatális csoporton belül, magasabb F0 estén, a két csoport között is jelentkezik. A diagramokon megfigyelhető még, hogy azokban az esetekben, mikor az F0 a beszédben megjelenő F1 értéke fölé kerül, az F1 az (elvártan megfelelően) valóban az alaphang frekvenciájára hangolódik.

Az [a:]–[i:]–[u:] hanghármas a magyar nyelv artikulációs bázisában a három egymástól legtávolabb eső hang, így F1-F2 értékeik függvényében ábrázolva meghatározzák nyelvünk magánhangzóterét. Ennek redukciója, tehát a hangok ejtésének centralizálódása figyelhető meg a 4. ábrán.

4. ábra

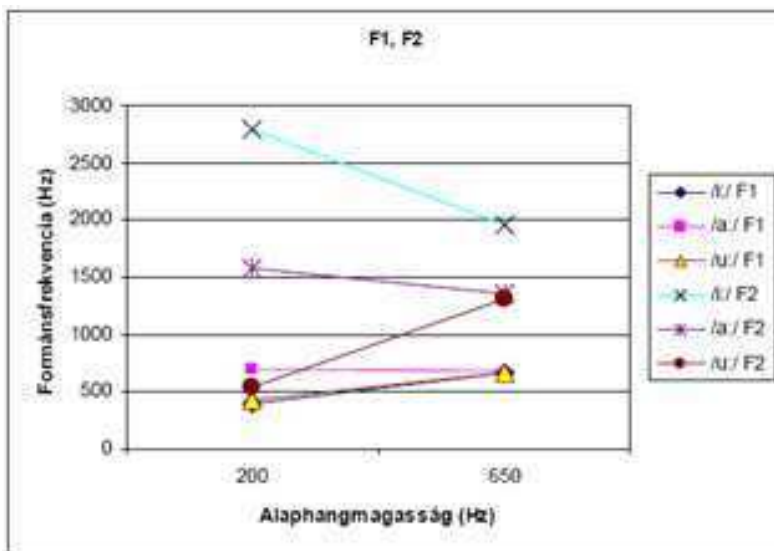
A magánhangzó-minőségek elkülönülése és az F0 emelése során centralizálódó ejtés (Függőleges vonal jelöli az F0 értékét, a hangokat jelölő pontokat összekötő élek pedig a magánhangzóteret határozzák meg.)



A centralizáció szempontjából kritikus alaphangmagasságot a kísérlet első fázisa során nem sikerült elérnünk, így a hangok teljes „összeolvadása” nem látható, de a tendencia így is érzékelhető. A redukálódott magánhangzóteret egyértelműen azt jelenti, hogy a nyelvünkre jellemző magánhangzó-minőségek elkülönülésének mértéke csökkent, aminek következtében perceptuális differenciálásuk is nehezebbé válik. A hangok formánsértékeinek változását szemlélteti az 5. ábra.

5. ábra

Az [a:], [i:] és [u:] hangok F1-F2 értékeinek változása

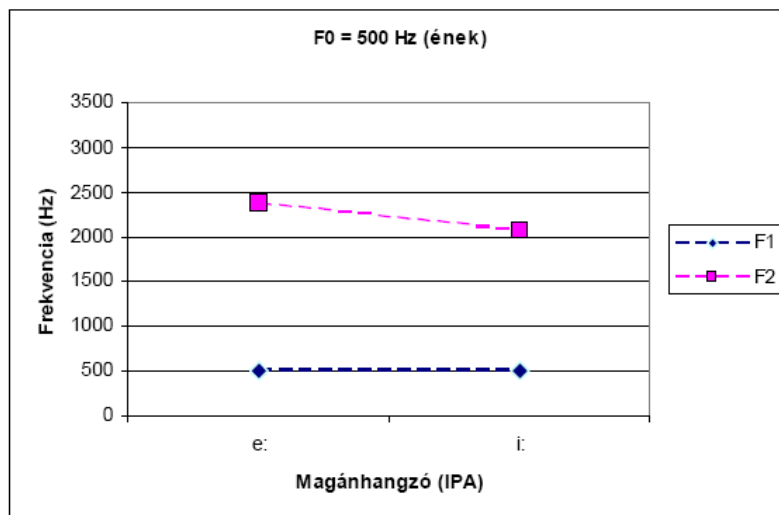
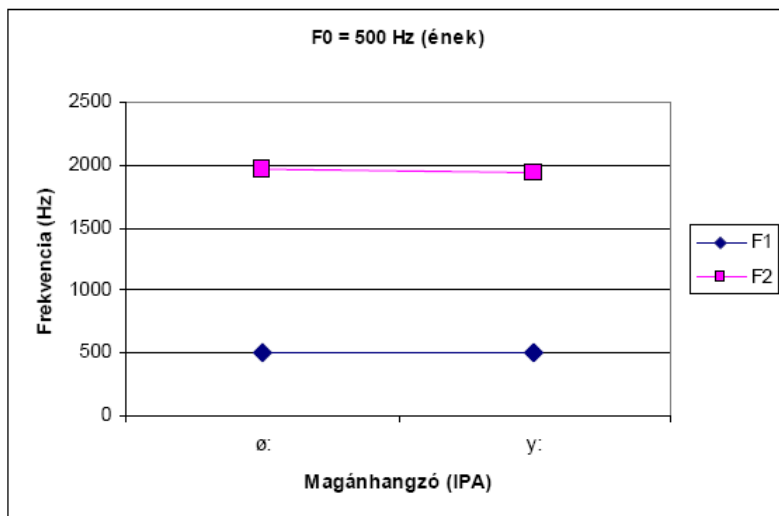


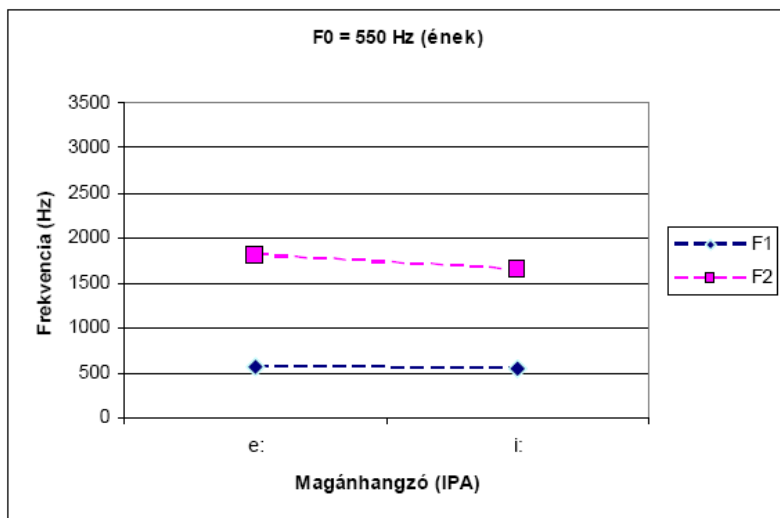
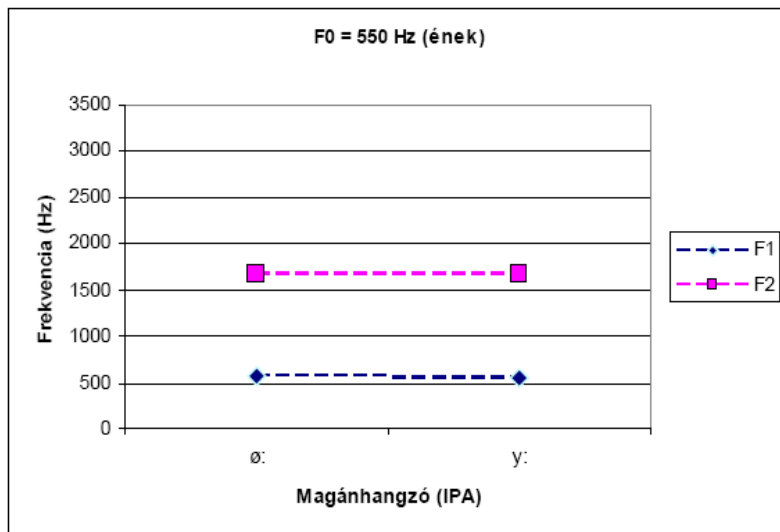
A legmagasabb alaphang mindhárom vokális első formánsának értékét saját frekvenciájára hangolta. A centralizált helyzetű vokális F1-F2 értékeire a vizsgált hangok megfelelő formánsértékeinek változási tendenciája utal.

A legmagasabb alaphangon is az ejtési szándéknak megfelelően legjobb arányban felismert három hang közül kettő ([i:], [y:]) jó eredményeit nem magyarázza maradéktalanul az artikuláció: a vokálisok képzéséhez a mindennapi beszédben működtetett artikulációs gesztusok és az éneklés artikulációs kényszere nem azonos irányba mutató tendenciák. E zárt állkapoccsal képzett hangok akusztikai szerkezetének változása jelentős: 1. az [y:] F1-F2 értékének aránya a beszédben 3:1 volt, míg 650 Hz-es alaphangmagasság mellett 2:1-re változott, 2. az [i:] F1-F2 értékeinek aránya a beszédben megjelenő 7:1-ről 2:1-re változott. A változás mindkét esetben arra utal, hogy jelentősen átalakult az eredeti hangszerkezet, így tehát a jó felismerési arányt az akusztikum stabilitása sem magyarázza. A következetesen tévesztett vokálisok vizsgálatára kapott eredményeket a 6. ábrában összesítettük.

6. ábra

A következőketen tévesztett magánhangzó-realizációk összevetése alaphang-magasságonként (folytatás a következő oldalon)





Az első két énekelt alaphangmagasságon a következetes tévesztésekben szereplő hangpárok tagjainak akusztikai szerkezete nagymértékben hasonlónak vált: F1 értékeik azonos, az F0 értékével megegyező frekvenciaértéken realizálódtak, F2 értékeik kicsi, maximum 300 Hz-es különbséggel jelentek meg. Az akusztikum ilyen mértékű hasonulása az oda-vissza tévesztést feltételeznél, ellenben ilyen csak az F0=550 Hz-es felvételen találunk az [ø:]–[y:] – [y:]–[ø:] esetében, ugyanis itt 50-50%-ban találkoztunk mindkét párral. A többi esetben nem lép fel nagy számban az ellenoldali tévesztés, ami tehát azt jelenti, nem egyértelmű, hogy az akusztikum hasonlósága indokolja a zártabb hangzó észlelését. A harmadik (F0=650 Hz) alaphangmagasságon anomáliát mutat az [ø:]–[i:] hangpár, itt ugyanis nagyobb a hangrealizációk F2 értékének eltérése (750 Hz), míg a percepció tesztelés azt is kimutatta, hogy az [ø:] ezen a magasságon egyáltalán nem felismerhető. A hangtévesztést az akusztikai hasonlóság nem magyarázhatja, az [ø:] ejtésének akadályozottságát viszont egyértelműen leolvashatjuk.

A kutatás során felmerült kérdés a hangidőtartamok és az észlelés közötti összefüggés. A jelen mérések alapján nem mutatható ki egyértelmű összefüggés a tiszta fázis időtartama és az ejtési szándéknak megfelelő észlelés között, hiszen a legmagasabb alaphangon az egyik legnagyobb arányban felismert [ɛ] realizációja volt a legrövidebb (39 ms), míg az egyik legkisebb arányban felismert [ɔ] realizációja a leghosszabb (78 ms). De ezt a kérdést további kutatásainkban tervezzük részletesebben is vizsgálni.

4. Következtetések. – Jelen kutatás az első magyar nyelven született, az éneklés artikulációs és akusztikai jellemzőinek a magyar nyelv artikulációs bázisára és vokális-realizációira gyakorolt hatását vizsgáló munka. A nemzetközi szakirodalom korábbi eredményeivel sajátjaink csak részben feleltethetők meg: 1. Igazolódott, hogy bizonyos mértékben megjelenik az éneklés artikulációs kényszerét követő tendencia, a nyíltabb hangzók zártabbként észlelése (ami valószínűsíti, hogy az ejtés is zártabban történik), de jellemző volt, hogy nagyobb arányban találkoztunk ennek fordított esetével, a zártabbként észleléssel. Minden alaphangmagasságon ezt mutatták a legnagyobb számban következetesen tévesztett hangok is. A nyílt magánhangzók elvárt jó felismerési aránya csak az [ɛ] hang esetében következett be, a két másik jó arányban pontosan észlelt hangzó az artikulációs kényszer csak részben (illabiális zárt [i:]), vagy egyáltalán nem segítő ejtési konfigurációjú (labiális zárt [y:]) vokális volt. 2. Az éneklés során az illabiális ejtésre törekvés az elvártan megfelelően megmutatkozik, a három legmagasabb F0 mellett legjobb arányban felismerhető magánhangzó közül kettő ([ɛ], [i:]) illabiális. A delabializálódással ellentétes tendenciát, azaz az illabiális helyett labiális ejtést és észlelést elvárásainknak megfelelően csak elenyésző számban találtuk meg a hibák között. 3. A hibák hierarchiájában az [a:]-ra tévesztés nagy aránya nem igazolódott. A hang pontos felismerése a legmagasabb alaphangon az elvárthoz képest jelentősen kisebb rátával teljesült (főként, hogy az elvárásainktól eltérően nem ez volt a legjobb arányban felismerhető vokális), a legtöbb tévesztés a zárt illabiális [i:]-re született. Ez az eredmény magyarázható azzal, hogy érvényesült az illabialitás mint az éneklés egyik fontos artikulációs kényszere, de kételyeket hagy afelől, hogy ho-

gyan lehet az, hogy az egyik legzártabban képzett hangzóról van szó. Az akusztikumra vonatkozó elvárások a percepciók teszteléséhez képest egyértelműbben teljesültek: 4. létrejött a formánskompresszió, és 5. egyértelműen be tudtuk bizonyítani a magánhangzótér redukcióját is. Az énekesformáns megjelenését nem állt szándékunkban igazolni, hiszen annak megjelenése női hangkategóriánál nem elvárt jelenség. Ellenőrző méréseink nem bizonyították a megjelenését.

A nemzetközi eredményekkel részben össze nem egyeztethető adatok magyarázataként egyrészt nyelvspecifikus jellemzők megjelenését feltételezzük: további kutatások igazolhatják, hogy az ejteni kívánt hangzók helyett nagyobb számban jelölt vokálisok észlelését a magyar nyelv magánhangzó-gyakorisága, illetve a magánhangzók realizációjának specifikus időtartama is befolyásolhatta. Másodsorban fontos lehet a hangkörnyezet hatása, hiszen jelen kutatásunk csak a nazális hangkörnyezetben ejtett magánhangzókra szorítkozott. Ebből következően a további kutatások egyik irányát képezheti más akusztikai kontextusok vizsgálata is. Harmadrészt: a nem egyenletes csökkenő felismerési arány megjelenésében szerepe lehet a vizsgált hangmagasságoknak is. Egy korábbi kísérlet (SCOTTO DI CARLO–GERMAIN 1985) leírta a regiszterváltás jelenségét, amely megmagyarázza az alaphangmagasság emelése és a hang ejtése közötti nem lineárisan változó összefüggéseket. Megfigyeléseik szerint a hangskála személyspecifikus regisztereinek felső határán a produkció romlik, a magánhangzó-minőségek kevésbé különíthetők el, míg a következő, magasabb regiszter alsó tartományának elérése, és ezáltal egy új fonációs pozíció létrehozása javít a produkción, így az (ejtési szándéknak megfelelő) azonosíthatóságon is. E jelenség vizsgálata szintén kutatási terveink között szerepel, ehhez szélesebb hangtartomány és kisebb F0-lépések rögzítése válik szükségessé. A negyedik, további kutatást igénylő terület a magánhangzótér teljes redukciójának elérése, és az ehhez szükséges akusztikai paraméterek (főként F0 értéke), illetőleg az így keletkező centralizálódott hang akusztikai paramétereinek (F1-F2 értékek) vizsgálata. Az ötödik, és jelen esetben egyben utolsó terület az énekesformáns jelensége. Az énekesformáns vizsgálatához férfi hangkategóriákkal terveztünk kísérletet, míg szándékunkban áll e nagy sáv szélességű intenzitás erősödés kimutathatóságának ellenőrzése (további akusztikai elemzési technikákkal) a női hangkategóriákban is. Mivel jelen kísérletünk résztvevője mindössze egy énekes és tíz hallgató személy volt, további kérdésekre ad választ a kutatás nagyobb számú kísérleti személlyel történő elvégzése is.

Az énekléstechnika produkciós és percepciók nehézségeinek deklarálása komoly következményekkel jár gyakorlati szempontból az énekes szakmára nézve is. Mivel a magas alaphangmagasság mellett a vokális artikulációs igénye sok esetben nem vág egybe az éneklés által megkívánttal, így egyértelművé válik, hogy a felismerhető beszéd generálásához inkább a mássalhangzók pontos ejtésének gyakorlása segítheti az énekest. Ennek megfelelően az énektanítás újragondolása is szükséges lehet. Másfelől az eredmények megnyugtatóak lehetnek az énekművészek számára is, hiszen az artikuláció nehézsége, mint láttuk, nem csak egyedi technikai megoldások, hanem univerzális artikulációs kényszerűségek következménye is. Ezen kívül pedig hangsúlyozható, hogy az operákban a közönség (anya)nyelvén tartott előadások feliratozása is kívánatos (ez Magyarországon jelenleg még nem, külföldön

sok helyütt már inkább jellemző), hiszen az énekelt beszéd értése, bizonyított módon, az anyanyelvünkön is nehézségekbe ütközik.

A hivatkozott irodalom

- GÓSY MÁRIA 1987. A formánsszerkezet változásának hatása a magánhangzók felismerésére. *Magyar Nyelv* 49–59.
- GÓSY MÁRIA 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- KERÉNYI MIKLÓS GYÖRGY 1959. *Az éneklés művészete és pedagógiája*. Zeneműkiadó, Budapest.
- KOVÁCS MAGDOLNA 1998. A spektrális minőség és az időtartam szerepe a magánhangzók percepciójában. *Beszédkutatás*, Budapest. 1998: 35–46.
- HOLLIEN, HARRY – MENDES-SWARTZ, ANA P. – NIELSEN, KENNETH 2000. Perceptual confusions of high-pitched sung vowels. *Journal of Voice*, Vol. 14., No. 2: 287–98.
- JOLIVEAU, ELOIDIE – SMITH, JOHN – WOLF, JOE 2004. Vocal Tract resonances in singing: The soprano voice. *Journal of Acousical Society of America*, Vol. 116, No. 4: 2434–9.
- LOSCOS, ALEX 2007. Spectral processing of the singing voice. Pompeu Fabra University, Barcelona. Doktori disszertáció. In: http://www.tesisenxarxa.net/TDX-0813107-163951/index_an.html. [Letöltve: 2009. 01. 30.]
- MILLHOUSE, THOMAS – CLERMONT, FRANTZ 2007. Acoustic description of a soprano's vowels based on percpetual linear prediction. 16th International Congress of Phonetic Sciences, Saarbrücken, 6–10 August, 2007: 901–4.
- MOLNÁR IMRE 1942. *Eufonétika. A szép beszéd és éneklés tana*. Kis Akadémia, Budapest.
- PETERSON, GORDON E. 1961. Parameters of vowel quality. *Journal of Speech and Hearing Research*, Vol. 4, No. 1: 10–29.
- SCOTTO DI CARLO, NICOLE – GERMAIN, ALINE 1985. A perceptual stuy of the influence of pitch on the intelligibility of sung vowels. *Phonetica* 42 (4): 188–97.
- SUNDBERG, JOHAN 1969. Articulatory differences between spoken and sung vowels in singers. *STL-QPSR*, Vol. 10, No. 1: 33–46.
- SUNDBERG, JOHAN 1970. Formant structure and articulation of spoken and sung vowels. *Folia Phoniatria*, Vol. 22, No. 1: 28–48.
- SUNDBERG, JOHAN 1972. An articulatory interpretation of the 'singing formant', *STL-QPSR*, Vol. 13, No. 1: 45–53.
- TARNÓCZY TAMÁS 1982. *Zenei akusztika*. Zeneműkiadó, Budapest.
- WENDAHL, RONALD R. 1959. Fundamental frequency and absolute vowel identification. *Journal of Acousical Society of America*, Vol. 31, No. 1: 109–10.

DEME ANDREA (andrea_deme@hotmail.com)

Phonetic analysis of Hungarian sung vowels

Features of articulation and perception of sung vowels at high fundamental frequencies are declared in several languages. Former studies have shown universal, and language-specific perceptual characteristics of vowels articulated with high fundamental frequency. This paper's aim is to investigate the acoustical, and perceptual effects of singing on Hungarian vowels. A female soprano singer

was asked to produce spoken and sung vowels in *mVn* structured nonsense-words. The recorded stimuli were perceptually tested and acoustically analysed in order to describe the perceptual and acoustical differences of spoken and sung vowels. Singing provides several articulatory differences compared to spoken speech, like lower jaw opening and drawing back the corner of the mouth (therefore illabial pronunciation). However, the acoustical measurements of the present study affirmed the vowel (space) reduction with F0-shift for Hungarian, some of the main tendencies of the perceptual experiment showed divergent tendencies of the vowel confusions. Results are presumably specific to the given language. Hence the results propose language-specific characteristics and let further questions arise.

DEME, ANDREA